Gene 427 (2008) 111–116

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Positive selection operates continuously on hemagglutinin during evolution of H3N2 human influenza A virus

Yoshiyuki Suzuki *

Center for Information Biology and DNA Data Bank of Japan, National Institute of Genetics, 1111 Yata, Mishima-shi, Shizuoka-ken 411-8540, Japan

ARTICLE INFO

Article history:
Received 2 July 2008
Received in revised form 8 September 2008
Accepted 10 September 2008
Available online 20 September 2008

Received by S. Yokoyama

Keywords:
H3N2 human influenza A virus
Hemagglutinin
Positive selection
Antigenic evolution

ABSTRACT

It has been proposed that antigenic evolution of hemagglutinin 1 (HA1) for H3N2 human influenza A virus was punctuated. In the population genetic analysis, however, it was controversial whether positive selection operated on HA1 in a punctuated manner for the branches of the phylogenetic tree where transitions to new antigenic clusters occurred (C branches), or continuously. In the molecular evolutionary analysis, positive selection was detected for the trunk (T) branches but the relationship between antigenic evolution and positive selection was unclear. Here molecular evolutionary analysis was conducted to examine natural selection operating on HA1 of H3N2 human influenza A virus by dividing HA1 into epitopes A–E and other sites, as well as dividing the phylogenetic tree into the C branches overlapping with the T branches (C–T branches), those not overlapping with the T branches (C–NT branches), the T branches not overlapping with the C branches (NC–T branches), and other branches (NC–NT branches). Positive selection was detected for C, T, and NC–T branches, whereas evolution for the NC–NT branches appeared to be mainly neutral. Positive selection appeared to have operated throughout the trunk, which covered the entire time period of the phylogenetic tree, suggesting that positive selection operated continuously on HA1 during evolution of H3N2 human influenza A virus.

1. Introduction

Influenza viruses are members of the family *Orthomyxoviridae* (Shope, 1931), containing a segmented and single-stranded (negative-sense) RNA genome in an enveloped virion (Noda et al., 2006). The genome encodes envelope glycoproteins, matrix proteins, nonstructural proteins, nucleoproteins, and polymerase proteins. Influenza viruses are classified into types A–C (Suzuki and Nei, 2002), among which the A virus is the most pathogenic to humans, causing 3–5 million cases of severe tracheobronchitis and 0.25–0.5 million deaths worldwide during annual epidemics (World Health Organization, 2003).

Hemagglutinin (HA) and neuraminidase (NA) are the envelope glycoproteins of influenza A virus. HA is cleaved into a signal peptide, HA1, and HA2 during maturation. HA1 is the major target of humoral immunity against influenza A virus. According to the antigenic properties of HA1, which are determined by epitopes A–E, and those of NA, influenza A viruses are classified into 16 (H1–H16) and 9 (N1–N9) subtypes, respectively (World Health Organization, 1980). H1N1 and H3N2 viruses are co-circulating in humans, among which the latter is more prevalent and pathogenic than the former (Nelson and Holmes, 2007).

When phylogenetic analysis was conducted for HA1s of H3N2 human influenza A viruses isolated in various years, it was found that isolates obtained in different years usually formed distinct clusters (Fitch et al., 1991). The clusters branched successively in the chronological order of isolation years from the trunk, which was defined as a set of branches [trunk (T) branches] connecting the root of the phylogenetic tree and the cluster of the latest isolates. It has been proposed that the trunk corresponds to evolution of the virus circulating continuously in the East and Southeast (E-SE) Asia or the tropics, and each cluster corresponds to annual epidemics in temperate regions, which are seeded from the E-SE Asia or the tropics and do not contribute to long-term evolution of the virus (Rambaut et al., 2008; Russell et al., 2008).

The antigenicity of influenza A virus evolves mainly by accumulation of amino acid substitutions (antigenic drift). When antigenic distances between HA1s of H3N2 human influenza A viruses isolated in various years were measured by the hemagglutination inhibition (HI) assay, isolates were found to form distinct antigenic clusters (Smith et al., 2004). It was proposed that antigenic evolution was

Abbreviations: HA(1, 2), hemagglutinin (1, 2); NA, neuraminidase; HI, hemagglutination inhibition; E-SE, East–Southeast; C branch, branch where a change to new antigenic cluster occurred; (N)T branch, (non-)trunk branch; (N)C–(N)T branch, (non-)trunk branch where a change to new antigenic cluster occurred (did not occur); NJ, neighbor-joining; TVM, transversional model; Γ, Γ distribution for the rate heterogeneity among sites; I, invariable sites; MP, maximum parsimony; r_S (r_N), rate of synonymous (nonsynonymous) substitution; c_S (c_N), number of synonymous (nonsynonymous) changes; s_S (s_N), number of synonymous (nonsynonymous) sites; p, probability.

* Fax: +81 55 981 6848.
E-mail address: yossuzuk@lab.nig.ac.jp.

0378-1119/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2008.09.012

punctuated, where the evolution was mostly small and occurred within antigenic clusters, and the cluster transition occurred only occasionally (de Jong et al., 2007). In contrast, evolution of nucleotide and amino acid sequences for HA1 appeared to be continuous.

To understand the mechanisms of antigenic evolution, it is interesting to detect natural selection operating on epitopes A–E of HA1. Natural selection operating on HA1 of H3N2 human influenza A virus has been examined by population genetic and molecular evolutionary analyses. In a population genetic analysis, a phylogenetic tree was constructed for HA1s of the viruses isolated in various years. For each T branch, all isolates were divided into antecedents and descendants, and the time interval required for extinction (fixation) of antecedents (descendants) was examined (Wolf et al., 2006). Some T branches were associated with rapid extinction of antecedents, whereas others with slow extinction, and it was proposed that evolution of HA1 was predominated by long-intervals of antigenic stasis, where neutral or nearly neutral amino acid substitutions accumulated to provide the basis of antigenic innovations, which were punctuated by short-intervals of antigenic changes, where positive selection operated on amino acid substitutions that caused antigenic changes through epistasis with the substitutions accumulated during the stasis (Koelle et al., 2006). However, frequencies of sequence variants for HA1 were also observed to change in nearly all the years under investigation, suggesting that positive selection was continuous (Shih et al., 2007).

In a molecular evolutionary analysis, the rates of synonymous (r_S) and nonsynonymous (r_N) substitutions were compared by dividing HA1 into epitopes A–E and other (non-epitope) sites, as well as dividing the phylogenetic tree into the T branches and other (NT) branches. It was observed that $r_N/r_S>1$ for epitopes A–E on the T branches, whereas $r_N/r_S<1$ for non-epitope sites and on the NT branches, suggesting that positive and negative selection operated for the former and the latter, respectively (Fitch et al., 1991; Wolf et al., 2006). However, the relationship between antigenic evolution and positive selection was unclear. The purpose of the present study was to examine this relationship by molecular evolutionary analysis.

2. Materials and methods

2.1. Sequence and antigenic data

The sequence and antigenic data of HA1s for 255 H3N2 human influenza A viruses isolated in various years were available in Smith et al. (2004). Since 89% of these isolates were propagated only in mammalian cell cultures for at most 5 passages (usually only 1 or 2 passages), the passage history did not appear to affect detection of

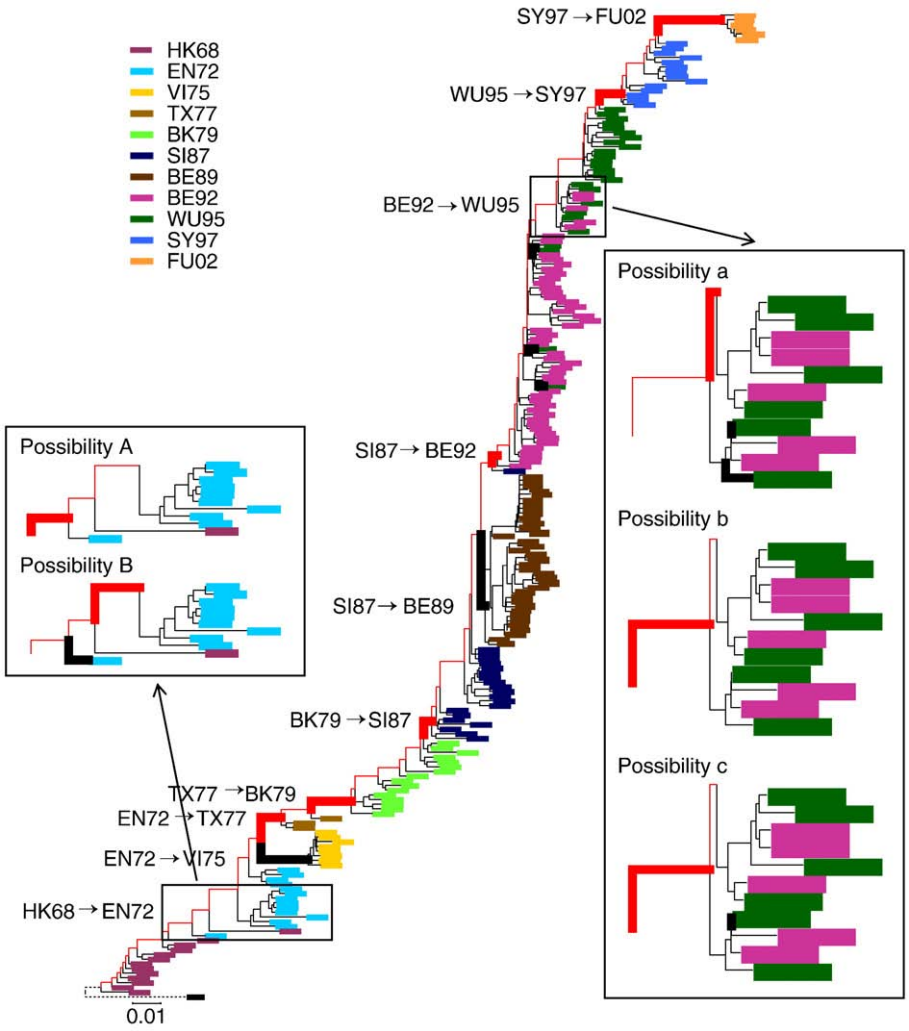

Fig. 1. Phylogenetic tree for HA1s of 209 H3N2 human influenza A viruses and 1 duck virus constructed by the NJ method assuming TVM+Γ. Each isolate is color-coded according to the antigenic cluster. The root of the human sequences is indicated as a broken line. The T branches are colored red, and the C branches are indicated as thick branches. Accordingly, the C–T, C–NT, NC–T, and NC–NT branches are indicated as thick red, thick black, thin red, and thin black branches, respectively. Two (possibilities A and B) and three (possibilities a, b, and c) equally parsimonious inferences of the C branches for the cluster transitions from HK68 to EN72 and from BE92 to WU95 are indicated in insets, respectively.

natural selection to any large extent (Bush et al., 2000; Zhai et al., 2007). After eliminating the isolates whose sequences contained ambiguous nucleotides, minor insertions, and minor deletions, as well as eliminating the isolates whose sequence and isolation year were both the same as others, 209 isolates were used for the analysis of natural selection. A duck sequence was added to identify the root of the phylogenetic tree for the human sequences. The sequences consisted of 987 nucleotide sites encoding the entire region of HA1 (329 amino acid sites). The human sequences have been classified into 11 antigenic clusters based on antigenic distances measured by the HI assay (Smith et al., 2004). The accession numbers in the International Nucleotide Sequence Database and antigenic clusters for the strains used in the present study are listed in Supplementary Table S1.

2.2. Data analysis

A multiple alignment for the total of 210 nucleotide sequences was made by using the computer program CLUSTAL W (version 1.83) (Thompson et al., 1994). The alignment did not contain any gaps. The phylogenetic tree for these sequences was constructed by the neighbor-joining (NJ) method (Saitou and Nei, 1987) assuming the model of nucleotide substitution that best fitted to the data. The model was identified by MODELTEST (Posada and Crandall, 1998) as the transversional model (TVM) with Γ distribution for the rate heterogeneity among sites [shape parameter (α) = 0.4959] (TVM + Γ) using the hierarchical likelihood-ratio test and TVM + Γ (α = 1.1365) with invariable sites (proportion of invariable sites = 0.2830) (TVM + Γ + I) using Akaike information criterion. PAUP* was used for these analyses (version 4.0b8a) (Swofford, 2000).

Natural selection was examined by dividing HA1 into epitopes A–E and non-epitope sites, as well as dividing the phylogenetic tree into the branches where transitions to new antigenic clusters (forward transitions) occurred (C branches), the T branches, and other (NC–NT) branches. The amino acid sites included in epitopes A–E are listed in Supplementary Table S2. The C branches were inferred by minimizing the total number of forward and backward (transitions to old antigenic clusters) transitions in the phylogenetic tree by the maximum parsimony (MP) method.

For the sites and branches under investigation, the total numbers of synonymous (c_S) and nonsynonymous (c_N) changes were counted based on the MP method (Suzuki and Gojobori, 1999). The average numbers of synonymous (s_S) and nonsynonymous (s_N) sites were computed by taking into account the transition/transversion rate ratio for the mutation pattern of H3N2 human influenza A virus, which has been estimated to be 7.5 in large-scale genomic analyses (Saitou, 1987; Suzuki, 2006). The null hypothesis of selective neutrality was tested by

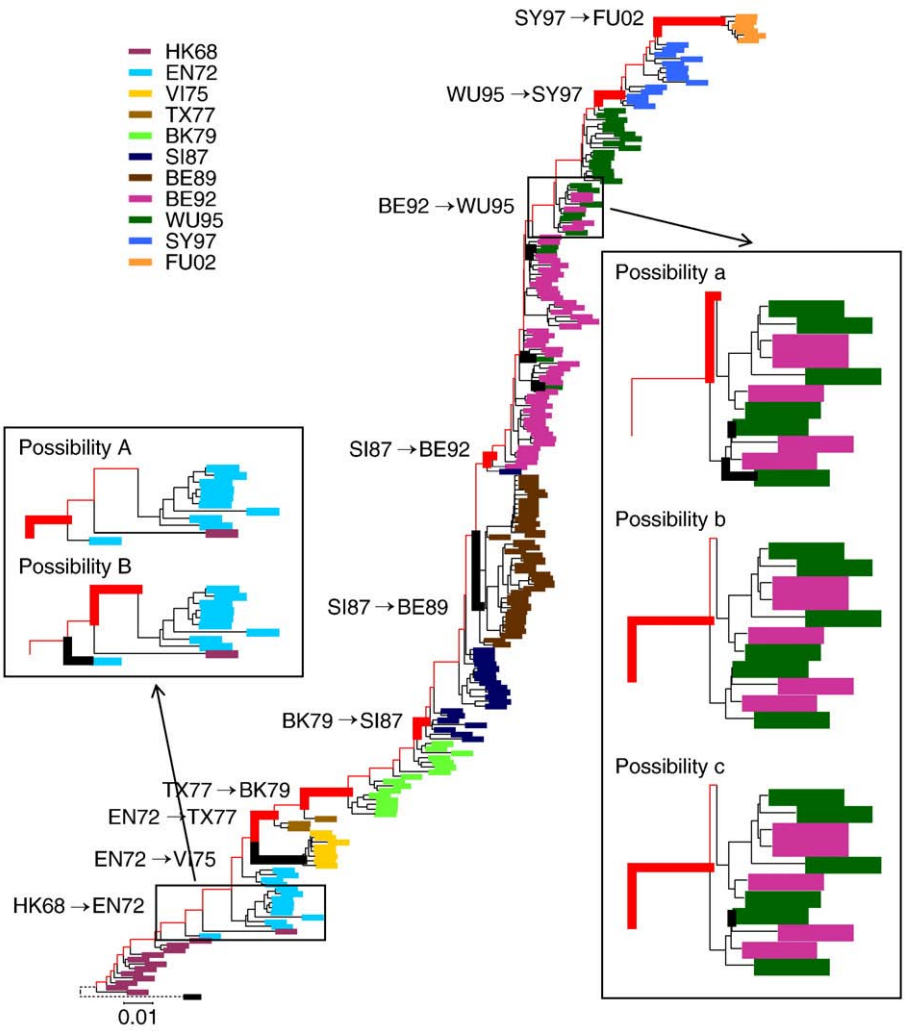

Fig. 2. Phylogenetic tree for HA1s of 209 H3N2 human influenza A viruses and 1 duck virus constructed by the NJ method assuming TVM + Γ + I. Each isolate is color-coded according to the antigenic cluster. The root of the human sequences is indicated as a broken line. The T branches are colored red, and the C branches are indicated as thick branches. Accordingly, the C–T, C–NT, NC–T, and NC–NT branches are indicated as thick red, thick black, thin red, and thin black branches, respectively. Two (possibilities A and B) and three (possibilities a, b, and c) equally parsimonious inferences of the C branches for the cluster transitions from HK68 to EN72 and from BE92 to WU95 are indicated in insets, respectively.

Table 1
r_N/r_S values estimated for various sites and branches for the phylogenetic tree constructed by the NJ method assuming TVM+Γ (Fig. 1)

Possibility	Epitope	Branches					
		C	T	C-NT	C-T	NC-T	NC-NT
A-a	A	**7.019**[a]	**5.192**	N.A.[b]	4.012	**6.361**	1.434
	B	**4.035**	2.367	N.A.	3.543	1.777	1.059
	C	0.503	0.606	0.209	0.700	0.587	0.692
	D	0.755	**0.490**	0.801	0.644	**0.481**	**0.532**
	E	0.776	1.662	N.A.	0.782	2.542	**0.602**
	Non	**0.147**	**0.066**	**0.101**	**0.163**	**0.037**	**0.132**
A-b	A	**6.811**	**5.192**	N.A.	4.407	5.962	1.471
	B	2.862	2.367	N.A.	2.535	2.198	1.059
	C	0.360	0.606	0.209	0.420	0.677	0.692
	D	0.777	**0.490**	0.803	0.752	**0.454**	**0.532**
	E	0.777	1.662	N.A.	0.784	2.538	**0.602**
	Non	**0.120**	**0.066**	**0.000**	**0.163**	**0.037**	**0.134**
A-c	A	**6.811**	**5.192**	N.A.	4.407	5.962	1.471
	B	2.862	2.367	N.A.	2.535	2.198	1.059
	C	0.360	0.606	0.209	0.420	0.677	0.692
	D	0.777	**0.490**	0.803	0.752	**0.454**	**0.532**
	E	0.777	1.662	N.A.	0.784	2.538	**0.602**
	Non	**0.120**	**0.066**	**0.000**	**0.163**	**0.037**	**0.134**
B-a	A	3.713	**5.192**	N.A.	2.003	**N.A.**	1.408
	B	**4.540**	2.367	3.005	5.059	1.467	1.060
	C	**0.210**	0.606	0.084	0.300	0.800	0.733
	D	0.808	**0.490**	0.941	0.430	**0.495**	**0.526**
	E	0.774	1.662	N.A.	0.781	2.543	**0.602**
	Non	**0.101**	**0.066**	**0.070**	**0.113**	**0.046**	**0.133**
B-b	A	3.609	**5.192**	N.A.	2.200	**N.A.**	1.446
	B	3.198	2.367	3.001	3.243	1.740	1.060
	C	**0.180**	0.606	0.084	0.233	0.978	0.733
	D	0.819	**0.490**	0.942	0.645	**0.469**	**0.526**
	E	0.776	1.662	N.A.	0.783	2.540	**0.602**
	Non	**0.081**	**0.066**	**0.000**	**0.113**	**0.046**	**0.135**
B-c	A	3.609	**5.192**	N.A.	2.200	**N.A.**	1.446
	B	3.198	2.367	3.001	3.243	1.740	1.060
	C	**0.180**	0.606	0.084	0.233	0.978	0.733
	D	0.819	**0.490**	0.942	0.645	**0.469**	**0.526**
	E	0.776	1.662	N.A.	0.783	2.540	**0.602**
	Non	**0.081**	**0.066**	**0.000**	**0.113**	**0.046**	**0.135**

[a] The r_N/r_S values are colored red and black when $r_N > r_S$ ($r_N/r_S > 1$) and $r_N \le r_S$ ($r_N/r_S \le 1$), respectively, and are bold-faced when positive (red) or negative (black) selection was detected.
[b] Not applicable because c_S (and r_S) was 0.

computing the probability (p) of obtaining the observed or more biased values for c_S and c_N under the assumption that these values followed a binomial distribution with the probabilities of occurrence of synonymous and nonsynonymous changes given by $s_S/(s_S+s_N)$ and $s_N/(s_S+s_N)$, respectively. Positive and negative selection were inferred when $c_N/s_N > c_S/s_S$ and $c_N/s_N < c_S/s_S$ with $p<0.05$, respectively. The r_N/r_S value was estimated as $(c_N/s_N)/(c_S/s_S)$. Although multiple changes were not corrected for obtaining c_S and c_N, the degree of underestimation appeared to be negligible because the branch lengths of the phylogenetic tree were generally small (Saitou, 1989).

3. Results

3.1. Phylogenetic tree for HA1

The phylogenetic trees for HA1s of 209 H3N2 human influenza A viruses and 1 duck virus constructed by the NJ method assuming TVM+Γ and TVM+Γ+I that were best fitted to the data are shown in Figs. 1 and 2, respectively. In both phylogenetic trees, each isolate is color-coded according to the antigenic cluster. The T branches, which were colored red in Figs. 1 and 2, were easily identified as a set of branches connecting the root of the human sequences and the cluster of the latest isolates. The C branches, which were indicated as thick branches in Figs. 1 and 2, were uniquely inferred for all cluster transitions, except for the transitions from HK68 to EN72 and from BE92 to WU95, where two (possibilities A and B) and three (possibilities a, b, and c) equally parsimonious inferences were made, respectively. Natural selection was examined under the assumption of each of six possibilities (A-a, A-b, A-c, B-a, B-b, and B-c) for the combinations of possibilities A and B and possibilities a, b, and c.

3.2. Natural selection operating on HA1

The r_N/r_S values estimated for various sites and branches under the assumption of possibility A-a for the phylogenetic tree constructed assuming TVM+Γ (Fig. 1) are summarized in Table 1. For the C branches, positive selection was detected for epitopes A and B. However, $r_N/r_S<1$ for epitopes C, D, E, and non-epitope sites, and negative selection was detected for the non-epitope sites.

It should be noted that positive selection has been detected for epitopes A–E on the T branches (Fitch et al., 1991). Indeed, for the T branches, $r_N/r_S>1$ for epitopes A, B, and E, and positive selection was detected for epitope A (Table 1). In the phylogenetic tree,

Table 2
r_N/r_S values estimated for various sites and branches for the phylogenetic tree constructed by the NJ method assuming TVM+Γ+I (Fig. 2)

Possibility	Epitope	Branches					
		C	T	C-NT	C-T	NC-T	NC-NT
A-a	A	**14.039**[a]	**5.191**	N.A.[b]	8.024	4.239	1.459
	B	**4.035**	2.367	N.A.	3.543	1.777	1.059
	C	0.503	0.606	0.209	0.700	0.587	0.692
	D	0.755	**0.489**	0.801	0.644	**0.481**	**0.540**
	E	0.776	1.564	N.A.	0.782	2.348	**0.602**
	Non	**0.151**	**0.071**	**0.101**	**0.170**	**0.041**	**0.133**
A-b	A	**13.624**	**5.191**	N.A.	8.813	3.973	1.496
	B	2.862	2.367	N.A.	2.535	2.198	1.059
	C	0.360	0.606	0.209	0.420	0.677	0.692
	D	0.777	**0.489**	0.803	0.752	**0.454**	**0.540**
	E	0.777	1.564	N.A.	0.784	2.344	**0.602**
	Non	**0.124**	**0.071**	**0.000**	**0.170**	**0.041**	**0.135**
A-c	A	**13.624**	**5.191**	N.A.	8.813	3.973	1.496
	B	2.862	2.367	N.A.	2.535	2.198	1.059
	C	0.360	0.606	0.209	0.420	0.677	0.692
	D	0.777	**0.489**	0.803	0.752	**0.454**	**0.540**
	E	0.777	1.564	N.A.	0.784	2.344	**0.602**
	Non	**0.124**	**0.071**	**0.000**	**0.170**	**0.041**	**0.135**
B-a	A	4.950	**5.191**	N.A.	2.670	**12.729**	1.433
	B	**4.539**	2.367	3.006	5.059	1.467	1.060
	C	**0.210**	0.606	0.084	0.300	0.801	0.733
	D	0.809	**0.489**	0.941	0.430	**0.495**	**0.534**
	E	0.775	1.564	N.A.	0.781	2.348	**0.602**
	Non	**0.103**	**0.071**	**0.070**	**0.117**	**0.051**	**0.134**
B-b	A	4.813	**5.191**	N.A.	2.934	11.932	1.471
	B	3.198	2.367	3.002	3.243	1.740	1.060
	C	**0.180**	0.606	0.084	0.233	0.979	0.733
	D	0.819	**0.489**	0.942	0.645	**0.469**	**0.534**
	E	0.776	1.564	N.A.	0.784	2.345	**0.602**
	Non	**0.083**	**0.071**	**0.000**	**0.117**	**0.051**	**0.136**
B-c	A	4.813	**5.191**	N.A.	2.934	11.932	1.471
	B	3.198	2.367	3.002	3.243	1.740	1.060
	C	**0.180**	0.606	0.084	0.233	0.979	0.733
	D	0.819	**0.489**	0.942	0.645	**0.469**	**0.534**
	E	0.776	1.564	N.A.	0.784	2.345	**0.602**
	Non	**0.083**	**0.071**	**0.000**	**0.117**	**0.051**	**0.136**

[a] The r_N/r_S values are colored red and black when $r_N > r_S$ ($r_N/r_S > 1$) and $r_N \le r_S$ ($r_N/r_S \le 1$), respectively, and are bold-faced when positive (red) or negative (black) selection was detected.
[b] Not applicable because c_S (and r_S) was 0.

however, some of the T branches overlapped with the C branches (Fig. 1), where positive selection was detected for epitopes A and B as indicated above. To examine whether positive selection operated for the C branches, T branches, or both, natural selection was examined by dividing the phylogenetic tree into the C branches overlapping with the T branches (C–T branches), those not overlapping with the T branches (C–NT branches), the T branches not overlapping with the C branches (NC–T branches), and the NC–NT branches. The C–T, C–NT, NC–T, and NC–NT branches are indicated as thick red, thick black, thin red, and thin black branches in Fig. 1, respectively. For the C–T, C–NT, and NC–T branches, the r_N value was always greater than the r_S value for epitopes A and B, and positive selection was detected for the NC–T branches. In contrast, for the NC–NT branches, $r_N/r_S \approx 1$ for epitopes A and B, and $r_N/r_S < 1$ and negative selection was sometimes detected for epitopes C, D, E, and non-epitope sites. Negative selection was detected for all the branches on the non-epitope sites.

The results obtained under the assumptions of possibilities A-b, A-c, B-a, B-b, and B-c were similar to those obtained under the assumption of possibility A-a, although the cases where positive and negative selection were detected were slightly different (Table 1). Similar results were also obtained when the phylogenetic tree was constructed assuming TVM+Γ+I (Fig. 2; Table 2).

4. Discussion

4.1. Positive selection operates continuously on HA1 during evolution of H3N2 human influenza A virus

It has been proposed that antigenic evolution of HA1 for H3N2 human influenza A virus was punctuated (Smith et al., 2004). In the population genetic analysis, however, it was controversial whether positive selection operated on HA1 in a punctuated manner for the C branches (Wolf et al., 2006), or continuously (Shih et al., 2007). In the molecular evolutionary analysis, positive selection was detected for the T branches (Fitch et al., 1991) but the relationship between antigenic evolution and positive selection was unclear. In the present study, by molecular evolutionary analysis, positive selection was detected for the NC–T branches, suggesting that antigenic changes were advantageous even when the magnitude of the change was relatively small. It should be noted that, for the C–T branches, where the magnitude of the change was relatively large, positive selection has been inferred in both of the population genetic analyses mentioned above (Wolf et al., 2006; Shih et al., 2007), and the r_N/r_S value was always >1 for epitopes A and B in the present study. Therefore, positive selection appears to have operated throughout the trunk (NC–T and C–T branches), which covers the entire time period of the phylogenetic tree. These observations suggest that positive selection operated continuously on HA1 during evolution of H3N2 human influenza A virus, although the positively selected sites may have changed on the C branches (Blackburne et al., 2008). These results were consistent with the observation that evolution of nucleotide and amino acid sequences for HA1 was continuous (Smith et al., 2004). In contrast, the r_N/r_S values for the NC–NT branches were ≈ 1 for epitopes A and B and <1 for epitopes C, D, E, and non-epitopes sites, where negative selection was sometimes detected, suggesting that evolution for the NC–NT branches was mainly neutral (Nelson et al., 2006).

4.2. Difference in natural selection operating on HA1 between geographical regions

It has been proposed that the trunk of the phylogenetic tree corresponds to evolution of H3N2 human influenza A virus in the E-SE Asia or the tropics, and each cluster corresponds to that in temperate regions (Rambaut et al., 2008; Russell et al., 2008). In the present study, positive selection was detected for the C, T, and NC–T branches, which were mostly included in or connected to the trunk, suggesting that positive selection mainly operated in the E-SE Asia or the tropics. In contrast, neutral evolution appeared to predominate for the NC–NT branches, suggesting that positive selection did not operate in temperate regions. The difference in natural selection between these geographical regions may be caused by the difference in the population dynamics of the virus. It has been reported that the transmission mode of influenza A virus may be different between the tropics and temperate regions, where contact or close-range spread appeared to play a major role in the former and aerosol transmission in the latter (Lowen et al., 2008). Since aerosol transmission appeared to be affected by temperature and relative humidity, whereas contact or close-range spread did not, influenza A virus may circulate continuously in the tropics but not in temperate regions. Therefore, natural selection may be expected to operate more efficiently in the tropics than in temperate regions, which was consistent with the results obtained in the present study. Since escape mutants appeared to be continuously generated in the E-SE Asia or the tropics, it may be important to monitor antigenic evolution of HA1 in these regions for predicting vaccine strains of H3N2 human influenza A virus.

Acknowledgments

The author thanks two anonymous reviewers for valuable comments. The present study was supported by KAKENHI 17770007 and KAKENHI 20570008 to Y.S.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.gene.2008.09.012.

References

Blackburne, B.P., Hay, A.J., Goldstein, R.A., 2008. Changing selective pressure during antigenic changes in human influenza H3. PLoS Pathog. 4, e1000058.
Bush, R.M., Smith, C.B., Cox, N.J., Fitch, W.M., 2000. Effects of passage history and sampling bias on phylogenetic reconstruction of human influenza A evolution. Proc. Natl. Acad. Sci. U. S. A. 97, 6974–6980.
de Jong, J.C., et al., 2007. Antigenic and genetic evolution of swine influenza A (H3N2) viruses in Europe. J. Virol. 81, 4315–4322.
Fitch, W.M., Leiter, J.M., Li, X.Q., Palese, P., 1991. Positive Darwinian evolution in human influenza A viruses. Proc. Natl. Acad. Sci. U. S. A. 88, 4270–4274.
Koelle, K., Cobey, S., Grenfell, B., Pascual, M., 2006. Epochal evolution shapes the phylodynamics of interpandemic influenza A (H3N2) in humans. Science 314, 1898–1903.
Lowen, A.C., Steel, J., Mubareka, S., Palese, P., 2008. High temperature (30 degrees C) blocks aerosol but not contact transmission of influenza virus. J. Virol. 82, 5650–5652.
Nelson, M.I., et al., 2006. Stochastic processes are key determinants of short-term evolution in influenza A virus. PLoS Pathog. 2, e125.
Nelson, M.I., Holmes, E.C., 2007. The evolution of epidemic influenza. Nat. Rev. Genet. 8, 196–205.
Noda, T., et al., 2006. Architecture of ribonucleoprotein complexes in influenza A virus particles. Nature 439, 490–492.
Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14, 817–818.
Rambaut, A., Pybus, O.G., Nelson, M.I., Viboud, C., Taubenberger, J.K., Holmes, E.C., 2008. The genomic and epidemiological dynamics of human influenza A virus. Nature 453, 615–619.
Russell, C.A., et al., 2008. The global circulation of seasonal influenza A (H3N2) viruses. Science 320, 340–346.
Saitou, N., 1987. Patterns of nucleotide substitutions in influenza A virus genes. Jpn. J. Genet. 62, 439–443.
Saitou, N., 1989. A theoretical study of the underestimation of branch lengths by the maximum parsimony principle. Sys. Zool. 38, 1–6.
Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425.
Shih, A.C.-C., Hsiao, T.-C., Ho, M.-S., Li, W.-H., 2007. Simultaneous amino acid substitutions at antigenic sites drive influenza A hemagglutinin evolution. Proc. Natl. Acad. Sci. U. S. A. 104, 6283–6288.
Shope, R.E., 1931. Swine influenza. III. Filtration experiments and etiology. J. Exp. Med. 54, 373–380.

Smith, D.J., et al., 2004. Mapping the antigenic and genetic evolution of influenza virus. Science 305, 371–376.
Suzuki, Y., 2006. Natural selection on the influenza virus genome. Mol. Biol. Evol. 23, 1902–1911.
Suzuki, Y., Gojobori, T., 1999. A method for detecting positive selection at single amino acid sites. Mol. Biol. Evol. 16, 1315–1328.
Suzuki, Y., Nei, M., 2002. Origin and evolution of influenza virus hemagglutinin genes. Mol. Biol. Evol. 19, 501–509.
Swofford, D.L., 2000. PAUP*: phylogenetic analysis using parsimony (and other methods). Sinauer, Sunderland, MA.
Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple-sequence alignment through sequence weighting, position-specific gap penalties, and weight-matrix choice. Nucleic Acids Res. 22, 4673–4680.
Wolf, Y.I., Viboud, C., Holmes, E.C., Koonin, E.V., Lipman, D.J., 2006. Long intervals of stasis punctuated by bursts of positive selection in the seasonal evolution of influenza A virus. Biol. Direct 1, 34.
World Health Organization, 1980. A revision of the system of nomenclature for influenza viruses: a WHO memorandum. Bull. WHO 58, 585–591.
World Health Organization, 2003. Influenza Fact Sheet. http://www.who.int/mediacentre/factsheets/fs211/en/index.html.
Zhai, W., Slatkin, M., Nielsen, R., 2007. Exploring variation in the d_N/d_S ratio among sites and lineages using mutational mappings: applications to the influenza virus. J. Mol. Evol. 65, 340–348.